Claims

1. An automated dialogue apparatus comprising:

- a buffer (10) for storing coded representations;

5 • speech generation means (6) operable to generate a speech signal from the coded representation for confirmation by a user;

- speech recognition means (2) operable to recognise speech received from the user and generate a coded representation of thereof;

- means (5) operable to compare the coded representation from the recogniser of a response

10 from the user with the contents of the buffer to determine, for each of a plurality of different alignments between the coded response and the buffer contents, a respective similarity measure, wherein at least some of said comparisons involve comparing only a leading portion of the coded response with a part of the buffer contents already uttered by the speech generation means; and

15 • means (5) for replacing at least part of the buffer contents with at least part of said recognised response, in accordance with the alignment having the similarity measure indicative of the greatest similarity.

2. An apparatus according to claim 1, including an input buffer operable to hold said coded
20 representation from the recogniser of a response from the user whilst said comparison is performed.

3. An apparatus according to claim 1, arranged so that said coded representation from the recogniser of a response from the user is entered into the buffer prior to said comparison , and
25 the replacing means is operable thereafter to adjust its position in the buffer.

4. An automated dialogue apparatus according to claim 1 or 2, further comprising means operable to divide the buffer contents into at least two portions, to supply an earlier portion to the speech generation means and to await a response from the user before supplying a later
30 portion to the speech generation means, wherein at least some of said comparisons involve comparing the coded response with a concatenation of a part of the buffer contents already

uttered by the speech generation means and the portion which, in the buffer, immediately follows it.

5. An apparatus according to any one of claims 1 to 4 including means operable to record status information defining the buffer contents as confirmed, offered for confirmation but not confirmed, and yet to be offered for confirmation.

6. An apparatus according to claim 5 in which the status information also includes indications of the condition that the respective coded representation has been corrected following non-confirmatory input from the user.

7. An apparatus according to claim 5 in which the status information is recorded by means of pointers indicating boundary positions within the buffer between representations having respective different status.

8. An apparatus according to claim 5 or 6 in which the buffer has a plurality of locations each for containing a coded representation, and for each location a status field for storing the associated status.

9. An apparatus according to claim any one of claims 5 to 8 in which the similarity measure is a function of (a) differences between the coded representation of the user's response and the contents of the buffer and (b) the status of those contents.

10. An apparatus according to any one of claims 5 to 9 in which the similarity measure is a function also of the alignment or otherwise of phrasal boundaries in the representations being compared.

11. An apparatus according to any one of claims 1 to 10 in which a portion of the coded representation of the user's response that in any particular alignment precedes the buffer contents is deemed to be different.

12. An apparatus according to any one of claims 1 to 11 in which a portion of the coded representation of the user's response that in any particular alignment follows the buffer contents does not contribute to the similarity measure.

13. An apparatus according to any one of claims 1 to 12 in which the comparing means is operable in accordance with a dynamic programming algorithm.

14. An apparatus according to any one of claims 1 to 12 in which the replacing means is operable, in the event that the alignment having the similarity measure indicative of the greatest similarity is an alignment corresponding to a pure continuation of the part of the buffer contents already uttered by the speech generation means, to enter the coded response into the buffer at such position and to mark the position within the buffer at which such entry began; and further comprising means operable to examine the buffer contents and to compare a part of the buffer contents immediately following a marked position with a part immediately preceding the same marked position to determine whether or not said immediately following part can be interpreted as a correction or partial correction of said immediately preceding part.

15. An apparatus according to claim 14 in which the replacing means is operable, in the event that the alignment having the similarity measure indicative of the greatest similarity is an alignment in which a non-leading portion of the coded response corresponds to a correction of that part of the buffer contents most recently uttered by the speech generation means, to insert the leading portion of the coded response into the buffer before the most recently uttered part, and to mark the position within the buffer at which such insertion began.

16. An apparatus according to claim 14 or claim 15, in which the means to examine and compare is operable in accordance with a dynamic programming algorithm.

17. An automated dialogue apparatus according to any one of claims 1 to 16, including means operable to recognise a spoken response containing an indication of non-confirmation and in response thereto to suppress selection of an alignment corresponding to a pure continuation of the part of the buffer contents already uttered by the speech generation means.

18. A method of speech recognition comprising

(a) receiving a coded representation;

(b) performing at least once the steps of

(b1) recognising speech from a speaker to generate a coded representation thereof;

5   (b2) updating the previous coded representation by concatenation of at least part thereof with this recognised coded representation;

(b3) marking the position within the updated representation at which said concatenation occurred; and

(c) comparing a part of the updated representation immediately following the marked position

10   with a part immediately preceding the same marked position to determine whether or not said immediately following part can be interpreted as a correction or partial correction of said immediately preceding part.

19. A method of speech recognition comprising

15   (a) recognising an utterance from a speaker to generate a coded representation thereof;

(b) detecting in the utterance a position that is followed by input having a correcting function and marking this position within the coded representation; and

(c) comparing a part of the updated representation immediately following the marked position with a part immediately preceding the same marked position to determine whether or not said

20   immediately following part can be interpreted as a correction or partial correction of said immediately preceding part.

20. A method according to claim 18 or 19 including performing the correction or partial correction.

25

21. A method according to claim 18 or 19 including performing the comparison in respect of a plurality of marked positions and performing the correction or partial correction in respect of that one of the marked positions for which a set criterion is satisfied.

30   22. A method according to claim 18 or 19 including performing the comparison in respect of a plurality of marked positions and performing the correction or partial correction in respect of a plurality of marked positions for which a set criterion is satisfied

23. A method according to claim 21 or 22 in which the set criterion is that the corrected

35   updated representation corresponds to an expected length.

24. A method according to claim 21 or 22 in which the set criterion is that the corrected updated representation matches a predetermined pattern definition.

25. A method according to any one of claims 18 to 24 including, in step (b), examining the recognised coded representation to determine whether it is to be immediately interpreted as a correction or partial correction, and performing such correction or partial correction, including continuation, if any;

wherein the steps of concatenation and marking are performed only in the event that the recognised coded representation is determined as not to be immediately interpreted as a correction or partial correction.

26. A method according to any one of claims 18 to 25 including generating, for confirmation, a speech signal from only part of the current coded representation, wherein said concatenation occurs at the end of that part.

27. A method according to any one of claims 18 to 26 in which the coded representation of step (a) is also generated by recognition of speech from the speaker.

28. A method according to any one of claims 18 to 27 in which:

step (b) is performed at least twice;

step (c) comprises performing a plurality of evaluations corresponding to different selections of one or more of said marked positions;

wherein each evaluation comprises performing said comparison in respect of the or each selected marked position and generating a cost measures as a function of the similarity determined by said comparison(s);

and wherein the question of which selection is to be chosen is determined based on said cost measure.

29. A method according to claim 28 in which said plurality of evaluations also include evaluations of the same selection of two or more marked positions processed in a different order.

30. A method according to any one of claims 18 to 29 in which said comparison is performed by means of a dynamic programming algorithm.

31. A method of speech recognition comprising

(a) recognising speech received from a speaker and generating a coded representation of each discrete utterance thereof; and storing a plurality of representations of discrete utterances in sequence in a buffer, including markers indicative of divisions between units corresponding to the discrete utterances;

(b) performing a comparison process having a plurality of comparison steps, wherein each comparison step comprises comparing a first comparison sequence (each of which comprises a unit or leading portion thereof) with a second comparison sequence which, in the stored sequence, immediately precedes the first comparison sequence, so as to determine whether the first and second comparison sequences meet a predetermined criterion of similarity;

(c) in the event that the comparison process identifies only one instance of first and second comparison sequences meeting the criterion, deleting the second comparison sequence of that instance from the stored sequence.

32. A method of speech recognition comprising

(a) recognising speech received from a speaker and generating a coded representation of each discrete utterance thereof; and storing a plurality of representations of discrete utterances in sequence in a buffer, including markers indicative of divisions between units corresponding to the discrete utterances;

in response to a parameter which defines an expected length for the stored sequence, the step of comparing the actual length of the stored sequence with the parameter and in the event that the actual length exceeds the parameter:

(b) performing a comparison process having a plurality of comparison steps, wherein each comparison step comprises comparing a first comparison sequence (each of which comprises a unit or leading portion thereof) with a second comparison sequence which, in the stored sequence, immediately precedes the first comparison sequence, so as to determine whether the first and second comparison sequences meet a predetermined criterion of similarity;

(c) in the event that the comparison process identifies only one instance where both (i) the length of the second comparison sequence is equal to the difference between the actual and

expected length and (ii) the first and second comparison sequences meet the criterion, deleting the second comparison sequence of that instance from the stored sequence.

33. A method according to claim 31 or 32 comprising, in the case that no deletion is
5   performed at step (c), performing a further such comparison process having a different predetermined criterion and/or a different manner of selection of the first and second comparison sequences.

34. A method of speech recognition comprising
10   (a) storing a coded representation;
     (b) selecting a portion of the stored coded representation;
     (c) supplying the selected portion to speech generation means operable to generate a speech signal therefrom for confirmation by a user;
     (d) recognising a spoken response from the user to generate a coded
15   representation thereof; and
     (e) updating the stored coded representation on the basis of the recognised response;
     wherein said updating includes updating at least one part of the stored coded representation other than the selected portion.

20   35. A method according to claim 34 including the step of (f) repeating steps (b) to (d) at least once.

36. A method according to claim 34 or 35 including generating for each selected portion a first marker indicative of the position thereof within the stored coded
25   representation.

37. A method according to any one of claims 34 to 36 in which said updating includes, according to the content of the recognised coded representation, one or more of:
     (i) correcting the selected portion or part thereof;
30   (ii) entering at least part of the recognised coded representation into the stored coded representation at a position immediately following the selected portion.

38. A method according to claim 37 in which said updating includes, according to the content of the recognised coded representation,

(iii) inserting a leading part of the recognised coded representation into the stored coded representation at a position before the selected portion.

5

39. A method according to claim 37 or 38 including generating for each entered part and any inserted part a second marker indicative of the position thereof within the stored coded representation.

10   40. A method according to claim 39 comprising the subsequent step of comparing, for the or each second marker, a part of the updated representation immediately following a position marked by that second marker with a part immediately preceding the same marked position to determine whether said immediately following part can be interpreted as a correction or partial correction of said immediately preceding part.

15

41. A method according to claim 40 when dependent on claim 25 in which said subsequent step of comparing compares a part of the updated representation immediately following a position marked by a second marker preferentially or exclusively with one or more immediately preceding parts marked by a first marker.

20

42. An automated dialogue apparatus comprising

speech generation means operable to generate a speech signal from a coded representation for confirmation by a user, characterised by means operable in dependence on the length of the coded representation to divide the coded representation into at least two portions, to supply a

25   first portion to the speech generation means and to await a response from the user before supplying any further portion to the speech generation means.

43. An apparatus according to claim 42 including means for recognising predetermined patterns in the coded representation and wherein upon such recognition one of the portions is

30   determined by reference to a recognised pattern.

44. An automated dialogue apparatus comprising:

- speech generation means operable to generate a speech signal from a coded representation for confirmation by a user; and

- means operable to divide the coded representation into at least two portions, to supply a first portion to the speech generation means and to await a response from the user before supplying any further portion to the speech generation means;

characterised by means for recognising predetermined patterns in the coded representation and wherein upon such recognition one of the portions is determined by reference to a recognised pattern.

45. An apparatus according to claim 43 or 44 in which the predetermined patterns are predetermined digit sequences occurring at the commencement of the representation.

46. An apparatus according to claim 45 for recognising telephone numbers, in which the coded representation is a representation of numeric digits.

47. An apparatus according to claim 45 or 46 in which the remainder of the coded representation is divided into portions such that each such portion shall not exceed a predetermined length.

48. An apparatus according to any one of claims 42 to 47 including speech recognition means operable to recognise speech received from the user and generate the coded representation therefrom.

49. An automated dialogue apparatus comprising:

- a buffer (10) for storing coded representations;
- speech recognition means (2) operable to recognise speech received from the user, including detecting phrasal boundaries in said input speech, and to store in the buffer a coded representation of the recognised speech and markers indicative of the positions of said phrasal boundaries;

- speech generation means (6) operable to generate a speech signal from the coded representation for confirmation by a user;
- control means operable in response to the phrase boundary markers to divide the coded representation into at least two portions, to supply a first portion to the speech generation
5 means for a response from the user before supplying any further portion to the speech generation means.

50. An automated dialogue method comprising:

storing coded representations including markers indicative of points of ambiguity;
10 comparing, for each of a plurality of different alignments thereof, a part of the coded representations immediately following a marked point with a part immediately preceding the same marked point to determine whether or not said immediately following part can be interpreted as a correction or partial correction of said immediately preceding part; wherein at least some of said comparisons involve comparing only a leading portion of said
15 immediately following part with said immediately preceding part.

51. An automated dialogue apparatus comprising:

- speech recognition means operable to recognise speech received from a speaker and
20 generate a coded representation thereof;

- timeout means operable to determine in accordance with a silence duration parameter when an utterance being recognised is deemed to have ended;

characterised by means operable, during an utterance, in dependence on the contents of the utterance to date, to vary the timeout parameter for the continuation of that utterance.

25

52. An automated dialogue apparatus according to claim 51 in which said variation is conditional upon the initial part of the utterance matching a predetermined pattern

53. An automated dialogue apparatus according to claim 51 in which said variation is
30 conditional upon recognition in the utterance of input indicative of negative confirmation to increase the timeout parameter for the remainder of that utterance.

54. An automated dialogue apparatus comprising:

- speech recognition means operable to recognise speech received from a speaker and generate a coded representation thereof;

- timeout means operable to determine in accordance with a silence duration parameter when an utterance being recognised is deemed to have ended;

characterised by means operable in dependence on a dialogue state to vary the timeout parameter.